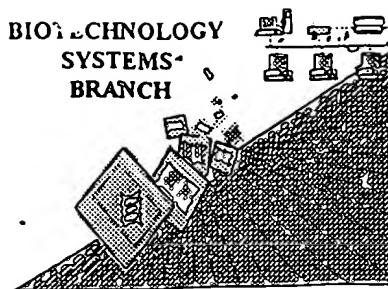


**RAW SEQUENCE LISTING**  
**ERROR REPORT**

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



970  
1119

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/863901

Source: OIPE

Date Processed by STIC: 11/16/2001

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:**

**Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER:

09/863901

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY P

- 1      Wrapped Nucleics  
    Wrapped Aminos  
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to 3; this will prevent "wrapping."
- 2      Invalid Line Length  
The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering  
The numbering under each 3<sup>rd</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII  
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length.  
Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"  
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)  
Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(ii) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)  
Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence  
<210> sequence id number  
<400> sequence id number  
000
- 9 ✓ Use of n's or Xaa's  
    (NEW RULES)  
Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents
- 10      Invalid <213>  
    Response  
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or Artificial Sequence
- 11      Use of <220>  
Sequence(s)      missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"  
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n  
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MII - Biotechnology Systems Branch - 08/21/2001

\* Various other errors and error types: See  
Examine sequence listing pages 1 and 2.  
error report attached. mtt

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/863,901

DATE: 11/16/2001

TIME: 09:16:49

Input Set : C:\CRF3\Datahold\Sample.app

Output Set: N:\CRF3\11162001\I863901.raw

3 <110> APPLICANT: Smith, John  
 4 Smith, Jane  
 5 Garcia Consulting  
 7 <120> TITLE OF INVENTION: Test Application Title (CCH00001)  
 9 <130> FILE REFERENCE: CCH01  
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/863,901  
 C--> 12 <141> CURRENT FILING DATE: 2001-05-24  
 14 <160> NUMBER OF SEQ ID NOS: 7  
 16 <170> SOFTWARE: PatentIn For Windows v. 2

**Does Not Comply**  
**Corrected Diskette Needed**

## ERRORED SEQUENCES

73 <210> SEQ ID NO: 2  
 74 <211> LENGTH: 950  
 75 <212> TYPE: DNA  
 76 <213> ORGANISM: Uca pugilator  
 78 <220> FEATURE:  
 79 <221> NAME/KEY: CDS  
 80 <222> LOCATION: (273)...(389)  
 81 <223> OTHER INFORMATION: Test Other Information  
 83 <400> SEQUENCE: 2  
 84 ccattttaaac caaaaaaaaa aaacggtgga caccggttaa ccccggttaa gtaccggttt 60  
 86 aggccatttc aggccaaatg tgcccaacta cgccaattgt ttgccaacg gccaacgtta 120  
 88 cgttcgtacg cacgtatgta cctaggtact tacggacgtg actacggaca cttccgtacg 180  
 90 tacgtacgtt tacgtaccca tcccaacgta accacagtgt ggtcgcagtg tcccagtgt 240  
 92 cacagactgc cagacattct tcacagacac cc cat gac acc acc tga acg tct 293  
 93 His Asp Thr Thr Thr Ser  
 W--> 94 1 5  
 96 ctt cct ccc aag ggt gtg tgg cac cac cct aca cct cct cct tct ggg 341  
 97 Leu Pro Pro Lys Gly Val Trp His His Pro Thr Pro Pro Ser Gly  
 W--> 98 10 15 20  
 100 gct gct gct ggt tct gct gcc tgg ggc cca tgt gag gca gca gga gaa 389  
 101 Ala Ala Ala Gly Ser Ala Ala Trp Gly Pro Cys Glu Ala Ala Gly Glu  
 W--> 102 25 30 35  
 104 tggggtggct cagccaaacc ttgagcccta gageccccct caactctgtt ctccctagggg 449  
 106 ctcatgcac ttgccacag caacctcaaa cctgctgtct acctcattgt aaacatccac 509  
 108 ctgacctccc agacatgtcc ccaccagctc tctctctacc cctgcctcag gaaccaagc 569  
 110 atccacctcc ctcccccaac ttcccccaag ctaaaaaaaa cagagggagc ccaactcctat 629  
 112 gctccccct gccatcccc aggaactcag tcagtgccca cttctacccc agcaagcaga 689  
 114 actcactgct ctggagagca aacacggacc gtgccttctt ccaggatggt ttctccttga 749  
 116 gcaacaattc tctcctggtc tagaaaaaat aattgatttc aagaccttct cccattctg 809  
 118 cctccattct gaccatttca ggggtcgtca ccacctctcc tttggccatt ccaacagctc 869  
 120 aagtcttccc tgatcaagtc accggagctt tcaaagaagg aattctaggc atcccagggg 929  
 E--> 122 acccacacct cctgaacca 949  
 135 <210> SEQ ID NO: 4  
 136 <211> LENGTH: 38

①

Misaligned amino numbering  
 numeric representation should align  
 directly under the first letter in  
 the abbreviation.

①

②

Errored

Input 950 field 211  
 Found 949 line 122

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/863,901

DATE: 11/16/2001

TIME: 09:16:49

Input Set : C:\CRF3\Datahold\Sample.app

Output Set: N:\CRF3\11162001\I863901.raw

137 <212> TYPE: PRT  
 138 <213> ORGANISM: Uca pugilator  
 140 <400> SEQUENCE: 4  
 141 Thr Ser Leu Pro Pro Lys Gly Val Trp His His Pro Thr Pro Pro Pro  
 142 1 5 10 15  
 144 Ser Gly Ala Ala Ala Gly Ser Ala Ala Trp Gly Pro Cys Glu Ala Ala  
 145 20 25 30

E--&gt; 147 Gly Glu

171 &lt;210&gt; SEQ ID NO: 6

172 &lt;211&gt; LENGTH: 7

173 &lt;212&gt; TYPE: PRT

174 &lt;213&gt; ORGANISM: Uca pugilator

176 &lt;400&gt; SEQUENCE: 6

E--&gt; 177 Xaa Glu Xaa Xaa Xaa Xaa Xaa

178 1 5

181 &lt;210&gt; SEQ ID NO: 7

182 &lt;211&gt; LENGTH: 12

183 &lt;212&gt; TYPE: PRT

184 &lt;213&gt; ORGANISM: Uca pugilator

186 &lt;400&gt; SEQUENCE: 7

E--&gt; 187 Xaa Xaa Xaa Xaa Xaa

E--&gt; 188 1 5

③ Errorred: must explain Xaa's in  
 fields 221, 222 and 223

④ Errorred: must explain Xaa's in  
 fields 221, 222 and 223

⑤ Sequence 4  
 Field 211 Input 38  
 Found 34 line 147

## VERIFICATION SUMMARY

DATE: 11/16/2001

PATENT APPLICATION: US/09/863,901

TIME: 09:16:50

Input Set : C:\CRF3\Datahold\Sample.app

Output Set: N:\CRF3\11162001\I863901.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:32 M:286 W: Invalid Database Entry: Valid Formats YYYY-MM-DD or Mon-YYYY, SEQ ID:1  
L:34 M:256 W: Invalid Numeric Header Field, Wrong PATENT FILING DATE:YYYY-MM-DD  
L:35 M:256 W: Invalid Numeric Header Field, Wrong PUBLICATION DATE:YYYY-MM-DD  
L:94 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:2  
L:98 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:2  
L:102 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:2  
L:122 M:252 E: No. of Seq. differs, <211>LENGTH:Input:950 Found:949 SEQ:2  
L:147 M:252 E: No. of Seq. differs, <211>LENGTH:Input:38 Found:34 SEQ:4  
L:163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:163 M:112 C: (48) String data converted to lower case,  
L:164 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:165 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5  
L:168 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:177 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:6  
L:187 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:7  
L:188 M:252 E: No. of Seq. differs, <211>LENGTH:Input:12 Found:5 SEQ:7